

# Object Recognition

Session-15

# Object Recognition (5-Questions)

# What are the key differences between object detection, identification, and categorization?

**Detection** locates objects (e.g., "Is there a person?").

**Identification** distinguishes a specific object (e.g., "Is that the Eiffel Tower?").

**Categorization** groups objects by type (e.g., "This is a building").

*In autonomous driving:  
detection finds a pedestrian,  
identification  
may recognize a traffic officer,  
and categorization could label all  
vehicles.*

# How many object categories can humans recognize intuitively?

- Humans can recognize thousands of object categories effortlessly due to years of experience and context understanding. Machines require extensive data and training for similar capability.



# In what ways does context influence object recognition?

- Context helps disambiguate objects (e.g., a “mouse” on a desk is likely a computer device, not an animal). In AI, incorporating scene context (e.g., “office” vs. “field”) improves accuracy.



# How is object recognition applied in autonomous vehicles like Tesla or Waymo?

## Used for

- pedestrian detection,
- traffic sign recognition,
- lane detection, and obstacle avoidance.

It's essential for decision-making and ensuring safety in dynamic environments.

# How would you design an object recognition system for a warehouse robot?

## Include

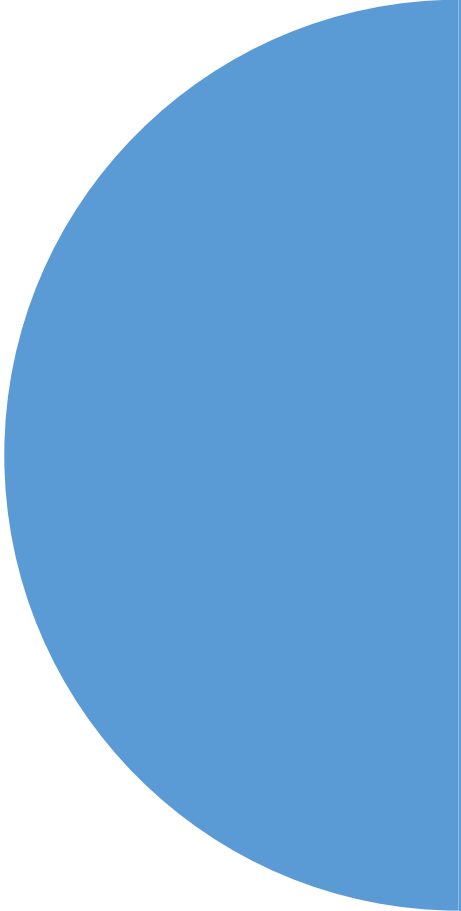
barcode  
reading,

item shape  
and color  
recognition,

object  
tracking for  
moving  
inventory, and

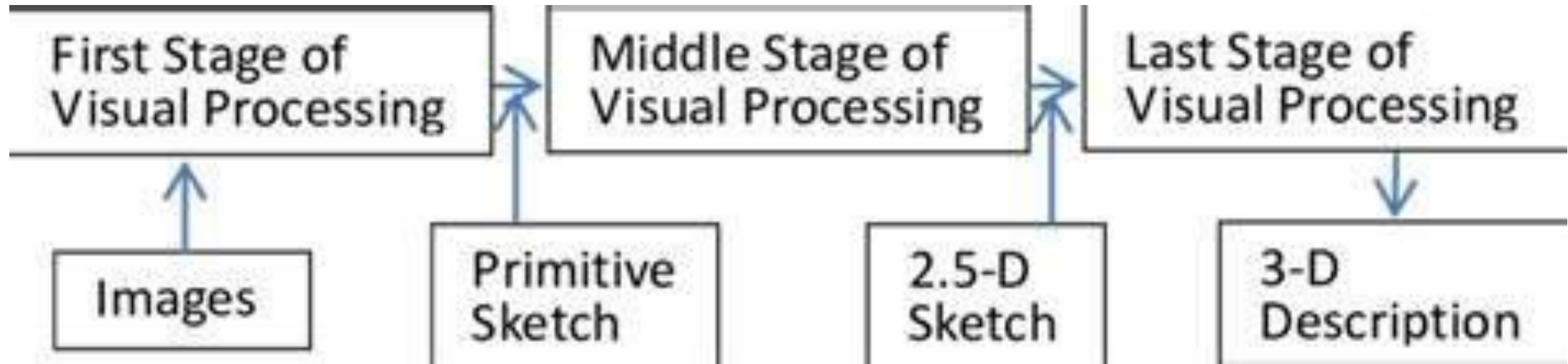
safety  
detection for  
human  
presence.

Should object recognition systems be allowed in public surveillance without consent? Why or why not?



Opinions vary. Supporters cite safety and crime prevention; critics cite privacy violations. A balanced approach includes transparency and opt-out mechanisms.

# Marr's Vision Paradigm



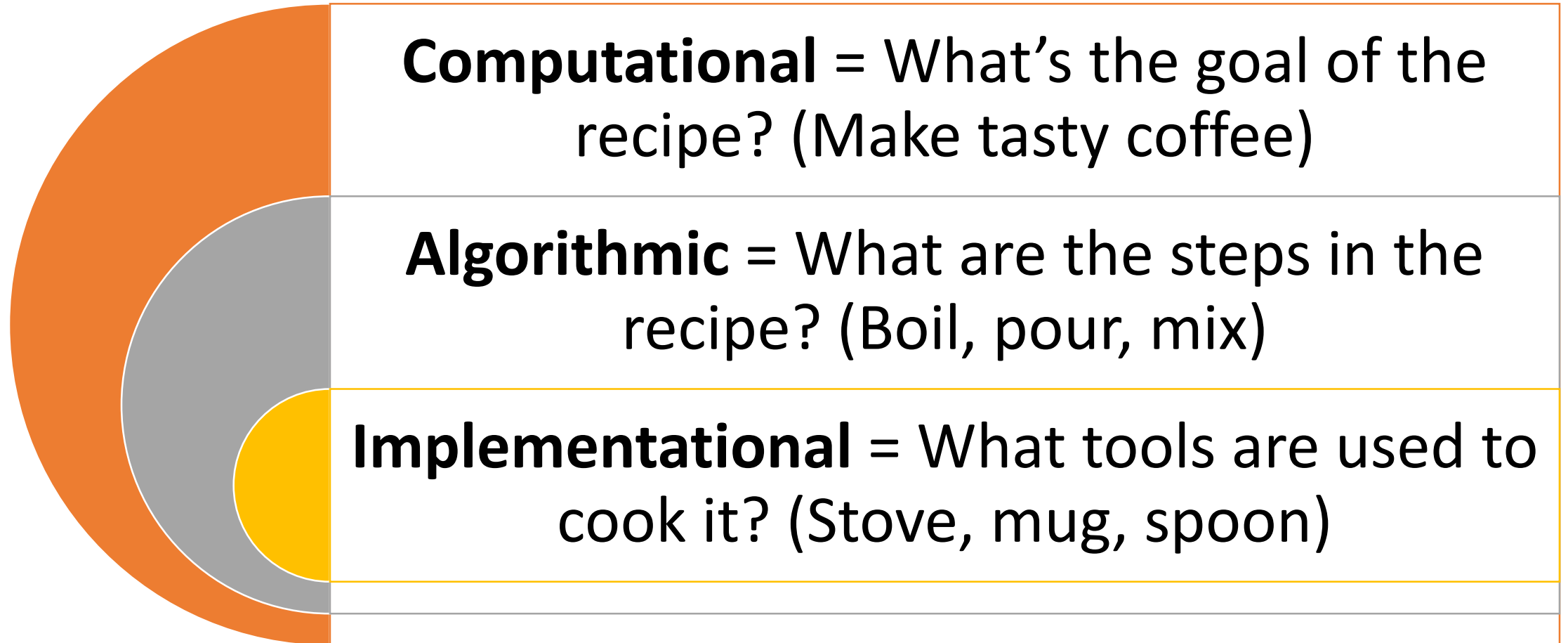


# Marr's Vision Paradigm as Making a Cup of Coffee with a Smart Robot

Imagine you have a **robot that makes coffee**. You want to understand **how it does it**. Marr would say you need to look at this robot on three levels:

- Computational level- What's the goal? What problem is being solved?
- Algorithmic level- How is it done? what steps & rules are used?
- Implementation level- what's it made of? How is it physically built to do this?

# 3-Levels of Marr's Vision Paradigm



# 3-Levels of Marr's Vision Paradigm

- Why does it make coffee?
  - What does “a good cup of coffee” mean?
  - This is about **what needs to be done**, not how.
- The robot:
  - Identifies a mug using vision,
  - Moves to the coffee machine,
  - Adds the right amount of water and coffee,
  - Presses buttons in the correct order.
  - This is the "**recipe**" or **procedure** it follows. It's like a **flowchart** of decisions and actions.
- Now we look at:
  - The robot's **camera** that sees the mug,
  - The **motors** in its arms,
  - The **circuit board** that runs the coffee algorithm.
  - It's about the **hardware** and physical stuff that carries out the plan.